

Using Householder Matrices to Establish Mixing Test Critical Values

Aaron Carl Smith
aaron.smith@ucf.edu

*Department of Mathematics
 University of Central Florida
 4000 Central Florida Blvd
 P.O. Box 161364
 Orlando, FL 32816-1364
 (407) 823-0538
 (407) 823-6253 (fax)*

Abstract: A measure-preserving dynamical system can be approximated by a Markov shift with a bistochastic matrix. This leads to using empirical stochastic matrices to measure and estimate properties of stirring protocols. Specifically, the second largest eigenvalue can be used to statistically decide if a stirring protocol is weak-mixing, ergodic, or nonergodic. Such hypothesis tests require appropriate probability distributions. In this paper, we propose using Monte Carlo empirical probability distributions from unistochastic matrices to establish critical values. These unistochastic matrices arise from randomly constructed Householder matrices.

AMS 2000 subject classifications: Primary 37A25, 62P30; secondary 37A05.

Keywords and phrases: weak-mixing, measure-preserving, stirring protocol.

Contents

1	Introduction	2
2	Bistochastic Matrices	4
3	Modeling Dynamical Systems	4
4	Mixing Hypothesis Test Procedure	6
5	Constructing the Monte Carlo Matrices	8
6	Establishing Critical Values	10
7	Matrix Convergence	13
8	Using Specific Random Variables	16

9 Two Region Partitions	21
References	22

1. Introduction

If a dynamical system has a probability measure and the system is measure-preserving, then partitioning the domain into n states of equal measure leads to a Markov shift whose measure is defined by a bistochastic matrix and the length n row vector

$$\left(\frac{1}{n} \dots \frac{1}{n}\right). \quad (1.1)$$

This partition approximation is called Ulam's method [21]. After partitioning the space, data point movement from one iteration of the function provides a stochastic matrix that approximates the bistochastic matrix from Ulam's method. Hence a Markov shift with an empirical stochastic matrix and the $1/n$ row vector approximates the dynamical system [see [17, Chapter 1, Chapter 9] for procedure and convergence rate].

We may model a stirring protocol's affect on a compression-resistant fluid with a measure-preserving dynamical system. In this paper we are interested in discrete iterations of a stirring protocol where the fluid at the beginning is the same fluid at the end. We 'look' at the fluid before and after stirring, but not during.

Properties of an empirical Markov shift can measure and evaluate properties of a measure-preserving dynamical system [8, 10, 13]. The second largest eigenvalue of a empirical stochastic matrix arising from Ulam's method may be used to statistically decide if a measure-preserving dynamical system is weak-mixing, ergodic, or nonergodic [6, 7, 9, 14, 17]. To statistically test if the dynamical system is ergodic, we need to have some knowledge of

$$P(|\widehat{\lambda}_2 - 1| > k : \lambda_2 = 1); \quad (1.2)$$

to statistically test if the dynamical system is weak-mixing, we need to have some knowledge of

$$P(|\widehat{\lambda}_2| > k : |\lambda_2| = 1). \quad (1.3)$$

We use λ_2 to denote the second largest eigenvalue of a bistochastic matrix arising from Ulam's method, and $\widehat{\lambda}_2$ denotes the second largest eigenvalue

of a corresponding empirical stochastic matrix (bistochastic and unistochastic matrices will be defined shortly). The utility of $\hat{\lambda}_2$ as a test statistic arises directly from the relationship between stochastic matrix eigenvalues and Markov shift ergodic, mixing properties. Since the unit circle contains all eigenvalues of stochastic matrices, there are no reasonable probability distributions of $\hat{\lambda}_2$ with 1 as the mean or median of either $\hat{\lambda}_2$ or $|\hat{\lambda}_2|$. So for hypothesis testing, we should use a probability distribution that has significant mass near $\hat{\lambda}_2 = 1$ or $|\hat{\lambda}_2| = 1$.

In this paper, we show that it is reasonable to approximate the conditional probability distributions with Monte Carlo probability distributions when the equal-measure partition sets are small. These Monte Carlo probability distributions are constructed using randomly generated Householder matrices.

Stirring protocols of compression resistant fluids, such as chocolate and water, provide examples of nearly measure-perserving dynamical systems. The need for confidence in the mixing of food items and in the mixing of pharmaceuticals highlights the utility of such probability distributions.

We propose using randomly generated, nonzero, independent, identically distributed real numbers to generate Householder matrices; take products of permutation matrices with Householder matrices; then square the magnitude of the products' entries to get unistochastic matrices. From these unistochastic matrices, construct a Monte Carlo approximation of a desired probability distribution. From the Monte Carlo probability distribution, establish the critical value for rejecting the null-hypothesis. The primary focus of this paper is to establish a method for determining hypothesis test critical values. Deciding which specific probability distribution to use in a hypothesis test depends on properties of the dynamical system; we will only show that the presented Monte Carlo methods are reasonable and leave probability distribution selection for the future.

There are several ways to use Monte Carlo methods to generate bistochastic matrices; unfortunately, many techniques lead to empirical probability distributions where the central tendency of $\hat{\lambda}_2$ is close to zero [17, Chapter 12]. Such distributions provide little utility for a weak-mixing, ergodic, or nonergodic hypothesis test. The methods presented here lead to Householder matrices that, in a Frobenius norm sense, are likely to be close to the identity matrix. Squaring the magnitude of entries from these unitary matrices gives unistochastic matrices. If we want a unistochastic matrix close to a particular

permutation matrix, we may multiply the Householder matrix by the desired permutation matrix. The main advantage of this method is that it provides probability distributions based on observed unistochastic matrices.

2. Bistochastic Matrices

Definition 2.1. *An $n \times n$ bistochastic matrix is a stochastic matrix whose transpose is also a stochastic matrix.*

By the Birkhoff-von Neumann theorem, the set of $n \times n$ bistochastic matrices form a convex set with permutation matrices as extreme points. We refer to this set as Birkhoff's polytope [2, 3]. Bistochastic matrices are also referred to as *doubly stochastic*.

Definition 2.2. *An $n \times n$ bistochastic matrix is called unistochastic if each entry is equal to the squared magnitude of some unitary matrix.*

The set of $n \times n$ unistochastic matrices form a proper subset of Birkhoff's polytope [2, page 307, section 1]. Since the set of unistochastic matrices is a proper subset, the proposed method should only be used when the Ulam's method bistochastic matrix is approximately unistochastic.

Definition 2.3. *An $n \times n$ Householder matrix is of the form*

$$H = I - 2\vec{v}\vec{v}^* \quad (2.1)$$

where \vec{v} is a unit vector.

Every Householder matrix is a unitary matrix [11, Chapter 5]. Since the set of $n \times n$ unitary matrices is closed under multiplication, taking the square magnitude of entries from a Householder matrix-permutation matrix product results in a unistochastic matrix.

3. Modeling Dynamical Systems

Consider running a stirring protocol on a compression resistant fluid. Let's model this with a measure-preserving dynamical system $(\mathbb{D}, \mathcal{B}, \mu, f)$,

1. \mathbb{D} represents the compression resistant fluid,
2. \mathcal{B} is the Borel σ -algebra,
3. μ is rescaled Lebesgue measure so that $\mu(\mathbb{D}) = 1$,
4. $f : \mathbb{D} \rightarrow \mathbb{D}$ models fluid movement during stirring.

The Monte Carlo method we will outline uses $n \times n$ unistochastic matrices arising from Householder matrix-permutation matrix products. These unistochastic matrices are close to the permutation matrices when n is large. It is reasonable to use the described Monte Carlo distribution for hypothesis testing when the dynamical system has the following property: For any $A, B \in \mathcal{B}$ where $P(A \cap B) = 0$, if f is perturbed so that

$$P(f(x) \in A \mid x \in B) \text{ increases (decreases),} \quad (3.1)$$

then for all Borel set $C \subseteq B^c$

$$P(f(x) \in A \mid x \in C) \text{ decreases (increases) proportionally.} \quad (3.2)$$

The $1/n$ row vector and unistochastic matrices arising from Householder matrix-permutation matrix products provide Markov shifts that reflect Ulam's method with an equal measure partition applied to such dynamical systems. This is not saying that all such dynamical systems lead unistochastic matrices, but that the Householder constructed unistochastic matrices reflect these properties.

If \vec{v} is a real unit vector and H is the corresponding Householder matrix,

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}, \quad H = I - 2\vec{v}\vec{v}^t. \quad (3.3)$$

Then squaring the entries of H gives a unistochastic matrix M ,

$$m_{ij} = \begin{cases} (1 - 2v_i^2)^2 & \text{if } i = j, \\ 4v_i^2 v_j^2 & \text{if } i \neq j. \end{cases} \quad (3.4)$$

If $\mathbb{D}_1, \mathbb{D}_1, \dots, \mathbb{D}_n$ are our equal measure partition sets, and M arose from Ulam's method, then the entries of M provide conditional probabilities,

$$m_{ij} = P(f(x) \in \mathbb{D}_j \mid x \in \mathbb{D}_i). \quad (3.5)$$

Increasing (decreasing) v_i^2 leads to nearly proportional decreases (increases) in

$$m_{ij} = P(f(x) \in \mathbb{D}_j \mid x \in \mathbb{D}_i) \text{ when } i \neq j. \quad (3.6)$$

Because of these observations, we propose using unistochastic Ulam matrices arising from squaring the entries of real Householder matrix-permutation matrix products to model weak-mixing stirring protocols of such dynamical systems.

4. Mixing Hypothesis Test Procedure

In this section we will discuss and outline our procedure for testing a compression-resistant fluid stirring protocol.

There are many techniques to measure a stirring protocol's ability to mix, such as decay of correlations [5], Fourier analysis, Artin braid patterns [1, 20], chaotic advection [18, 19], and other topological methods [4, 12]. Unfortunately, it is typical for different protocols to lend themselves to different analytical methods. Thus comparing mixing quality between mechanically dissimilar stirring protocols is difficult. The primary advantage of the Ulam method approximation is that it can be used to evaluate any incompressible fluid stirring pattern. This allows one to compare and evaluate the mixing of stirring protocols by comparing and evaluating eigenvalues. The main disadvantage is that the method is statistical and does not prove the results. Another significant advantage of our method is that it requires only one iteration of stirring, in contrast to other techniques that call for iterated experiments.

Since our method only approximates the dynamical system, we make no inferences regarding strong-mixing when we conclude that the stirring protocol is weak-mixing. If the protocol is not ergodic, then it is not weak-mixing. If the protocol is not weak-mixing, then it is not strong-mixing.

Our test hypotheses are

1. $H_o : (\mathbb{D}, \mathcal{B}, \mu, f)$ is not ergodic (and hence not weak-mixing).
2. $H_{a_1} : (\mathbb{D}, \mathcal{B}, \mu, f)$ is ergodic but not weak-mixing.
3. $H_{a_2} : (\mathbb{D}, \mathcal{B}, \mu, f)$ is weak-mixing (and hence ergodic).

We partition the fluid into connected, equal volume regions, and use these partition sets to generate a new σ -algebra contained in \mathcal{B} . If our data strongly indicate that the stirring protocol is weak-mixing or ergodic over the generated σ -algebra, we will conclude the same about the original dynamical system. If the stirring protocol is nonmixing or nonergodic over the generated σ -algebra, then the original dynamical system is nonmixing or nonergodic. The procedure evaluates stirring over a smaller σ -algebra, thus the test is inherently more reliable for detecting if a protocol is nonmixing or nonergodic.

The null hypothesis is that the stirring protocol is nonergodic. It is better to reject a protocol that mixes well than to produce poorly mixed product. The repercussions of a testing error are as follows:

1. Type I Error: discard a desirable stirring protocol for a different stirring

method

2. Type II Error: produce a product that is insufficiently mixed

The stirring protocol's purpose determines the tolerable risks of error and the number of partition states. If poorly mixed fluid could result in minor consequences or mixing on a small scale is inapt, then the number of partition regions may be relatively small. If poorly mixed fluid could result in severe consequences, then the number of partition regions must be large and partition volume small. For example, poorly mixed batter from a kitchen could result in unpalatable food; poorly mixed pharmaceuticals with a low LD50 could lead to overdose and death. A mixing test for a kitchen could use a relatively coarse partition, while a pharmaceutical company would use a fine partition.

If we know an upper bound for the stirring protocol's entropy, call it h , then Froyland's entropy estimate and expected values show that the number of states should be greater than e^h [8].

Data point movement from one iteration of stirring leads to our empirical stochastic matrix, \hat{P} . The percent of points that start in region i and end in region j gives us \hat{p}_{ij} . We model the entries of \hat{P} as nonindependent binomial random variables, whose probabilities come from the Ulam stochastic matrix. Our test statistic is $\hat{\lambda}_2 = \lambda_2(\hat{P})$. Some dynamical systems have measure zero sets with atypical properties. In an attempt to avoid such difficulties, we randomly select data points rather than select points from a grid.

If the data points are independent, uniform, and randomly distributed within each region, then the empirical matrix will converge to a bistochastic matrix in a Frobenius norm sense (the proof of this follows from extending a standard Monte Carlo argument [16]). We approximate the stirring protocol with a one-sided Markov shift, $\left(\left(\frac{1}{n}, \dots, \frac{1}{n}\right), \hat{P}\right)$. The pair will not define a Markov shift if

$$\left(\frac{1}{n}, \dots, \frac{1}{n}\right) \hat{P} \neq \left(\frac{1}{n}, \dots, \frac{1}{n}\right). \quad (4.1)$$

The $1/n$ row vector is a stationary distribution for any bistochastic matrix.

Since Ulam method's partitions our fluid into n equal volume regions, it is reasonable to use the $1/n$ vector as the stationary distribution. If we do not use equal volume partitions, the stationary distribution will be the probability vector corresponding to the rescaled volume of each region. Many of the results regarding convergence, expected values, convergence rates, etc.

depend on equal measure partitions, we should use equal volume regions if appropriate [17].

Mixing Hypothesis Test Procedure:

1. Set the type II error significance levels for both alternative hypotheses, α_1 and α_2 .
2. Set n to be the number of partition regions.
3. Decide which conditional probability distribution(s) for $|\lambda_2(P)| = 1$ and $\lambda_2(P) = 1$ to use.
4. Establish the critical values for H_{a_1} and H_{a_2} , c_1 and c_2 . The purpose of this paper is to propose using Householder matrix-permutation matrix products to estimate c_1 and c_2 .
5. Partition the fluid into n connected equal volume regions, $\mathbb{D}_1, \mathbb{D}_2, \dots, \mathbb{D}_n$.
6. Randomly select data points in each partition region. These points should be independent and uniformly distributed.
7. Run the stirring protocol one time.
8. Use data point movement between regions to construct an empirical stochastic matrix, \hat{P} .
9. Determine the hypothesis test result. The test statistic is $\lambda_2(\hat{P})$; compare $|\lambda_2(\hat{P}) - 1|$ to c_1 ; compare $|\lambda_2(\hat{P})|$ to c_2 .
10. Use Froyland's entropy estimate to estimate the dynamical system's entropy.

$$-\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n \hat{p}_{ij} \log \hat{p}_{ij} \quad (4.2)$$

(we define $0 \log 0$ to equal 0).

11. If the null hypothesis is rejected in favor of weak-mixing, let the rate at which

$$\binom{N}{n-1} (\lambda_2(\hat{P}))^{N-n+1} \rightarrow 0 \text{ as } N \rightarrow \infty \quad (4.3)$$

be our estimate of the rate of mixing.

5. Constructing the Monte Carlo Matrices

Let

$$n \in \{3, 4, 5, 6, 7, \dots\} \quad (5.1)$$

be the number of equal measure states that we partition the measure-preserving dynamical system into while using Ulam's method [see [17, Chapter 1] for the procedure]. Let

$$\{u_1, u_2, \dots, u_n\} \quad (5.2)$$

be real independent, identically distributed random variables such that

$$u_i \neq 0 \text{ almost surely, and } E\left(\frac{1}{u_i^8}\right), E(u_i^8), E\left(\frac{1}{u_i^4}\right), E(u_i^4) < \infty. \quad (5.3)$$

Let

$$\vec{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}, \vec{v} = \frac{\vec{u}}{|\vec{u}|}. \quad (5.4)$$

We may use a unit vector to construct a Householder matrix. Let $H = (h_{ij})$ be the Householder matrix corresponding to \vec{v} .

$$H = I - 2\vec{v}\vec{v}^T \quad (5.5)$$

The entries of H are

$$h_{ij} = \begin{cases} 1 - \frac{2}{u_1^2 + u_2^2 + \dots + u_n^2} u_i^2 & \text{if } i = j, \\ -\frac{2}{u_1^2 + u_2^2 + \dots + u_n^2} u_i u_j & \text{if } i \neq j. \end{cases} \quad (5.6)$$

Let Q be a permutation matrix that we want our random unistochastic matrix to be proximal to. Set $U = QH$. Now, let $M = (m_{ij})$ be the matrix defined by

$$m_{ij} = u_{ij}^2. \quad (5.7)$$

Since H is a Householder matrix and Q is a permutation matrix, U is a unitary matrix. It follows that M is unistochastic.

Notice that

$$m_{ij} = \begin{cases} \left(1 - \frac{2}{u_1^2 + u_2^2 + \dots + u_n^2} u_i^2\right)^2 & \text{if } i = j, \\ \frac{4}{(u_1^2 + u_2^2 + \dots + u_n^2)^2} u_i^2 u_j^2 & \text{if } i \neq j. \end{cases} \quad (5.8)$$

Since $u_i \neq 0$ almost surely for all i , all entries of M are positive almost surely; by the Perron-Frobenius theorem, all but one of M 's eigenvalues are of magnitude strictly less than one [15, Chapter 8]. It follows that any Markov shift with stationary distribution

$$\left(\frac{1}{n} \dots \frac{1}{n}\right) \quad (5.9)$$

will be strong-mixing (for Markov shifts, weak-mixing is equivalent to strong-mixing). Our hypothesis test for weak-mixing (ergodic) requires a probability distribution over $[0, 1]$ (the unit circle) with significant mass near 1. In the next two sections, we will see that the expected value of M 's eigenvalues converge to one as n goes to infinity.

How does our dynamical system relate to n ? Generally speaking, finer partitions are more apt to detect nonmixing (nonergodicity). If we are confident in mixing, we will use a coarse partition to reduce effort; if our confidence in mixing is poor, we will use a fine partition.

By the Birkhoff-von Neumann theorem, bistochastic matrices are convex combinations of permutation matrices [2, 3]. So our unistochastic matrices will tend to be near the 'corners' of the set of bistochastic matrices. For any statistic from unistochastic matrices we are interested in, we may use such matrices to generate a Monte Carlo empirical probability distribution.

6. Establishing Critical Values

In this section, we will outline the procedure we propose for establishing critical values for a weak-mixing, ergodic, nonergodic hypothesis test.

Our test hypotheses are

1. $H_o : (\mathbb{D}, \mathcal{B}, \mu, f)$ is not ergodic (and hence not weak-mixing).
2. $H_{a_1} : (\mathbb{D}, \mathcal{B}, \mu, f)$ is ergodic but not weak-mixing.
3. $H_{a_2} : (\mathbb{D}, \mathcal{B}, \mu, f)$ is weak-mixing (and hence ergodic).

After partitioning the space into n equal measure connected subsets, Ulam's method approximates the dynamical system with a Markov shift. We will approximate the bistochastic matrix defining the Markov shift's measure, P , with an empirical stochastic matrix, \hat{P} . So we approximate

$$(\mathbb{D}, \mathcal{B}, \mu, f) \text{ with } \left(\left(\frac{1}{n}, \dots, \frac{1}{n}\right), \hat{P}\right). \quad (6.1)$$

Remark 6.1. *The pair*

$$\left(\left(\frac{1}{n}, \dots, \frac{1}{n}\right), \hat{P}\right) \quad (6.2)$$

will not define a Markov shift if

$$\left(\frac{1}{n}, \dots, \frac{1}{n}\right) \widehat{P} \neq \left(\frac{1}{n}, \dots, \frac{1}{n}\right), \quad (6.3)$$

but if our data points are uniform random variables within each state, then

$$E(\|P - \widehat{P}\|_F) \rightarrow 0 \quad (6.4)$$

as the minimum number of points in a state goes towards infinity. It follows that for each eigenvalue

$$|\lambda_i(P) - \lambda_i(\widehat{P})| \rightarrow 0 \text{ in probability } \forall i \quad (6.5)$$

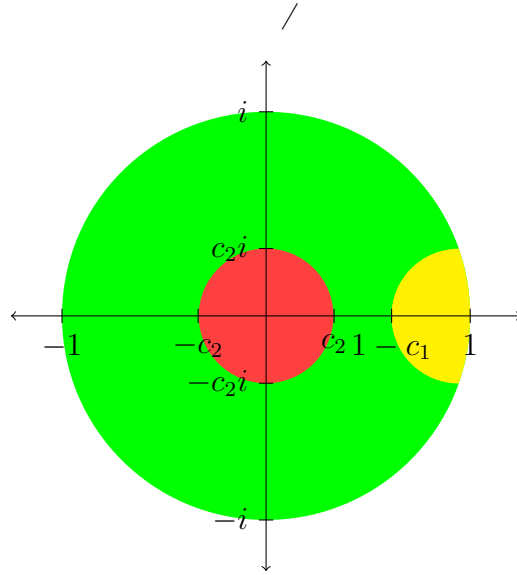
in the Hausdorff topology when our data points are uniform random variables within each state and the minimum number of points in a state goes towards infinity [17, Chapter 8].

Our test statistic is the second largest eigenvalue of \widehat{P} . Let $\alpha_1, \alpha_2 \in (0, 1)$ be the alpha values for the hypothesis test; let $c_1, c_2 \in (0, 1)$ be the corresponding critical values, $c_2 \leq 1 - c_1$,

$$P(|\lambda_2(\widehat{P}) - 1| \geq c_1 : \lambda_2(P) = 1) < \alpha_1, \quad (6.6)$$

$$P(|\lambda_2(\widehat{P})| \leq c_2 : |\lambda_2(P)| = 1) < \alpha_2. \quad (6.7)$$

Our goal is to use Householder matrices to estimate c_1 and c_2 . The probability distribution used to establish c_1, c_2 should reflect properties of a class of dynamical systems containing our stirring protocol.



Argand diagram of mixing hypothesis test criteria:

1. If $\lambda_2(\hat{P})$ is in the region containing 1 (yellow), fail to reject the null hypothesis; conclude that the dynamical system is nonergodic.
2. If $\lambda_2(\hat{P})$ is in the outer region away from 1 (green), reject the null hypothesis in favor of the first alternative hypothesis; conclude that the dynamical system is ergodic, but not weak-mixing.
3. If $\lambda_2(\hat{P})$ is in the center region (red), reject the null hypothesis in favor of the second alternative hypothesis; conclude that the dynamical system is weak-mixing (and hence ergodic).

Establishing Critical Values for the Test:

1. Partition \mathbb{D} into n equal measure connected subsets. If an upper bound of the dynamical system's entropy is known, call the upper bound h , set n greater than e^h [8].
2. Select a random variable with which to construct unit vectors.
Let u_1, u_2, \dots, u_n be independent, identically distributed, random variables,
 $\vec{v}_i = \frac{\vec{u}}{\|\vec{u}\|_2}$.
3. Set $N \in \mathbb{N}$ so that our empirical probability distributions will be sufficiently accurate.
4. Select permutation matrices $\{Q_i\}_{i=1}^N$ near which we want the probability distribution to have significant mass.
5. Randomly generate N Householder matrices, $H_i = I - 2\vec{v}_i\vec{v}_i^T$. Then

square the entries of $Q_i H_i$ to get the matrix M_i .

6. Use $\{\lambda_2(M_i)\}_{i=1}^N$ to approximate $P(|\lambda_2(\hat{P}) - 1| \geq k : \lambda_2(P) = 1)$. Use the approximation to estimate c_1 .
7. Use $\{|\lambda_2(M_i)|\}_{i=1}^N$ to approximate $P(|\lambda_2(\hat{P})| < k : \lambda_2(P) = 1)$. Use the approximation to estimate c_2 .

7. Matrix Convergence

In this section, we will show that M_i from the previous section will converge to the permutation matrix Q_i as n increases. Since permutation matrix eigenvalues are on the unit circle, as a random variable, it is likely that the second largest eigenvalue from one of our unistochastic matrices will be near magnitude one. Because of the likely proximity to one, it is reasonable to use a probability distribution from such an eigenvalue to establish critical values for our weak-mixing, ergodic, nonergodic hypothesis test.

Our proofs take advantage of the Frobenius norm. After using Jensen's inequality to remove the square root from consideration, finding an expected value upper bound is similar to finding a second moment. A permutation matrix acting on a matrix does not change the magnitude of the entries, without loss of generality will prove the results for when Q is the identity matrix and focus on M 's convergence to the identity matrix.

Proposition 7.1. *If M is a matrix constructed in section 2 with $Q = I$, $n \in \{3, 4, 5, \dots\}$, $\{u_i\}_{i=1}^n$ are identically distributed, $u_i \neq 0$ a.s. and*

$$E(u_i^4), E(u_i^8), E\left(\frac{1}{u_i^4}\right), E\left(\frac{1}{u_i^8}\right) < \infty, \quad (7.1)$$

then $E(\|M - I\|_F) \rightarrow 0$ as $n \rightarrow \infty$. Moreover,

$$E(\|M - I\|_F^2) \leq \frac{16n}{(n-1)^4} E\left(\frac{1}{u_i^8}\right) E(u_i^8) \quad (7.2)$$

$$+ \frac{16n}{(n-1)^2} E\left(\frac{1}{u_i^4}\right) E(u_i^4) \quad (7.3)$$

$$+ \frac{16n(n-1)}{(n-2)^4} E\left(\frac{1}{u_i^8}\right) (E(u_i^4))^2. \quad (7.4)$$

$$(7.5)$$

Proof. First, by Jensen's inequality

$$E(\|M - I\|_F) \leq \sqrt{E(\|M - I\|_F^2)}. \quad (7.6)$$

So it is sufficient to show the second part of the proposition. Let's look at the entries of $M - I$; by computation we see that:

$$(M - I)_{ij} = \begin{cases} -4\left(\frac{1}{u_1^2 + u_2^2 + \dots + u_n^2} u_i^2\right) \\ \quad \times \left(1 - \frac{1}{u_1^2 + u_2^2 + \dots + u_n^2} u_i^2\right) & \text{if } i = j, \\ \frac{4}{(u_1^2 + u_2^2 + \dots + u_n^2)^2} u_i^2 u_j^2 & \text{if } i \neq j. \end{cases} \quad (7.7)$$

It follows that

$$(M - I)_{ij}^2 = \begin{cases} \frac{16}{(u_1^2 + u_2^2 + \dots + u_n^2)^4} u_i^8 \\ \quad - \frac{32}{(u_1^2 + u_2^2 + \dots + u_n^2)^3} u_i^6 \\ \quad + \frac{16}{(u_1^2 + u_2^2 + \dots + u_n^2)^2} u_i^4 & \text{if } i = j, \\ \frac{16}{(u_1^2 + u_2^2 + \dots + u_n^2)^4} u_i^4 u_j^4 & \text{if } i \neq j. \end{cases} \quad (7.8)$$

If we expand the addends and remove the negative terms, it follows that almost surely

$$\|M - I\|_F^2 < \sum_{i=1}^n \frac{16}{(u_1^2 + u_2^2 + \dots + u_n^2)^4} u_i^8 \quad (7.9)$$

$$+ \sum_{i=1}^n \frac{16}{(u_1^2 + u_2^2 + \dots + u_n^2)^2} u_i^4 \quad (7.10)$$

$$+ \sum_{i \neq j} \frac{16}{(u_1^2 + u_2^2 + \dots + u_n^2)^4} u_i^4 u_j^4. \quad (7.11)$$

Almost surely, all of the terms in the denominators are positive; if we subtract terms from the denominators, we get an upper bound on the fractions. Thus almost surely

$$\|M - I\|_F^2 < \sum_{i=1}^n \frac{16}{(u_1^2 + u_2^2 + \dots + u_n^2 - u_i^2)^4} u_i^8 \quad (7.12)$$

$$+ \sum_{i=1}^n \frac{16}{(u_1^2 + u_2^2 + \dots + u_n^2 - u_i^2)^2} u_i^4 \quad (7.13)$$

$$+ \sum_{i \neq j} \frac{16}{(u_1^2 + u_2^2 + \dots + u_n^2 - u_i^2 - u_j^2)^4} u_i^4 u_j^4. \quad (7.14)$$

Notice that the numerator and denominator in each fraction in the upper bound are independent.

The subtraction of terms in denominators removed some positive terms, so the denominators are sums of positive terms. Therefore, we may use the harmonic-arithmetic means inequality,

$$\| M - I \|_F^2 < 16 \sum_{i=1}^n \frac{1}{(n-1)^8} \left(\frac{1}{u_1^2} + \frac{1}{u_2^2} + \dots + \frac{1}{u_n^2} - \frac{1}{u_i^2} \right)^4 u_i^8 \quad (7.15)$$

$$+ 16 \sum_{i=1}^n \frac{1}{(n-1)^4} \left(\frac{1}{u_1^2} + \frac{1}{u_2^2} + \dots + \frac{1}{u_n^2} - \frac{1}{u_i^2} \right)^2 u_i^4 \quad (7.16)$$

$$+ 16 \sum_{i \neq j} \frac{1}{(n-2)^8} \left(\frac{1}{u_1^2} + \frac{1}{u_2^2} + \dots + \frac{1}{u_n^2} - \frac{1}{u_i^2} - \frac{1}{u_j^2} \right)^4 u_i^4 u_j^4 \quad (7.17)$$

Now let's take expected values; since the u_i 's are independent,

$$\frac{E(\| M - I \|_F^2)}{16} \leq \sum_{i=1}^n E \left(\left(\frac{1}{u_1^2} + \frac{1}{u_2^2} + \dots + \frac{1}{u_n^2} - \frac{1}{u_i^2} \right)^4 \right) \frac{E(u_i^8)}{(n-1)^8} \quad (7.18)$$

$$+ \sum_{i=1}^n E \left(\left(\frac{1}{u_1^2} + \frac{1}{u_2^2} + \dots + \frac{1}{u_n^2} - \frac{1}{u_i^2} \right)^2 \right) \frac{E(u_i^4)}{(n-1)^4} \quad (7.19)$$

$$+ \sum_{i \neq j} E \left(\left(\frac{1}{u_1^2} + \frac{1}{u_2^2} + \dots + \frac{1}{u_n^2} - \frac{1}{u_i^2} - \frac{1}{u_j^2} \right)^4 \right) \frac{E(u_i^4)E(u_j^4)}{(n-2)^8}. \quad (7.20)$$

Next we use Minkowski's inequality and the fact that the u_i 's are independently distributed,

$$E(\| M - I \|_F^2) \leq \frac{16n}{(n-1)^4} E\left(\frac{1}{u_i^8}\right) E(u_i^8) \quad (7.21)$$

$$+ \frac{16n}{(n-1)^2} E\left(\frac{1}{u_i^4}\right) E(u_i^4) \quad (7.22)$$

$$+ \frac{16n(n-1)}{(n-2)^4} E\left(\frac{1}{u_i^8}\right) (E(u_i^4))^2. \quad (7.23)$$

$$(7.24)$$

Since $E(\frac{1}{u_i^8})$, $E(u_i^8)$, $E(\frac{1}{u_i^4})$, and $E(u_i^4)$ are all finite, it follows that

$$E(\| M - I \|_F^2) \rightarrow 0 \quad (7.25)$$

as $n \rightarrow \infty$. Hence by Jensen's inequality,

$$E(\| M - I \|_F) \rightarrow 0 \quad (7.26)$$

as $n \rightarrow \infty$. □

Since the second largest eigenvalue gives a test statistic to decide if a measure-preserving dynamical system is weak mixing, ergodic, or nonergodic, we need a conditional probability distribution of eigenvalues to conduct hypothesis tests. To statistically test if a measure-preserving dynamical system is weak-mixing, we could randomly select permutation matrices

$$\{Q_k\}_{k=1}^N \quad (7.27)$$

and generate $\{M_k\}_{k=1}^N$, with our Householder method, then use the empirical probability distribution from

$$\{|\lambda_2(M_k)|\}_{i=1}^N \quad (7.28)$$

to establish the critical value for the weak-mixing hypothesis test.

To statistically test if a measure-preserving dynamical system is ergodic, we could randomly select permutation matrices

$$\{Q_k : \text{the multiplicity of } \lambda = 1 \text{ is at least two}\}_{k=1}^N \quad (7.29)$$

and use Householder matrices to generate $\{M_k\}_{k=1}^N$, then use the empirical probability distribution from

$$\{\lambda_2(M_k)\}_{k=1}^N \quad (7.30)$$

to establish a critical value for the ergodic hypothesis test.

8. Using Specific Random Variables

In this section, we find more precise upper bounds for specific random variables. These upper bounds give better estimates of convergence rate than the results in the previous section. The first two proofs in this section start out the same way as the first proof in the previous section, then the arguments take advantage of the distribution properties.

Let's find a more precise upper bound when the u_i 's are independent standard normal random variables. The proof is similar to the first convergence proof, the difference is that we take advantage of the relationship between normal random variables and χ^2 -distributions.

Proposition 8.1. *If the u_i 's in the construction of a unistochastic matrix are independent standard normal random variables and $n \in \{11, 12, 13, \dots\}$, then*

$$E(\| M - I \|_F^2) < \frac{1680n}{(n-3)(n-5)(n-7)(n-9)} \quad (8.1)$$

$$+ \frac{48n}{(n-3)(n-5)} \quad (8.2)$$

$$+ \frac{144n(n-1)}{(n-4)(n-6)(n-8)(n-10)}. \quad (8.3)$$

Proof. From the previous proof, we know that almost surely

$$\| M - I \|_F^2 < \sum_{i=1}^n \frac{16}{(u_1^2 + u_2^2 + \dots + u_n^2 - u_i^2)^4} u_i^8 \quad (8.4)$$

$$+ \sum_{i=1}^n \frac{16}{(u_1^2 + u_2^2 + \dots + u_n^2 - u_i^2)^2} u_i^4 \quad (8.5)$$

$$+ \sum_{i \neq j} \frac{16}{(u_1^2 + u_2^2 + \dots + u_n^2 - u_i^2 - u_j^2)^4} u_i^4 u_j^4. \quad (8.6)$$

Since u_i 's are independent standard normal random variables, we may replace the u_i 's with χ^2 -random variables when we take expected values.

$$E(\| M - I \|_F^2) < \sum_{i=1}^n 16E\left(\frac{1}{(\gamma_{n-1})^4}\right)E(u_i^8) \quad (8.7)$$

$$+ \sum_{i=1}^n 16E\left(\frac{1}{(\gamma_{n-1})^2}\right)E(u_i^4) \quad (8.8)$$

$$+ \sum_{i \neq j} 16E\left(\frac{1}{(\gamma_{n-2})^4}\right)E(u_i^4)E(u_j^4). \quad (8.9)$$

We use γ_j to denote a χ^2 -random variable with j degrees of freedom. If we take expected values and remove negative terms, it follows that

$$E(\| M - I \|_F^2) < \frac{1680n}{(n-3)(n-5)(n-7)(n-9)} \quad (8.10)$$

$$+ \frac{48n}{(n-3)(n-5)} \quad (8.11)$$

$$+ \frac{144n(n-1)}{(n-4)(n-6)(n-8)(n-10)}. \quad (8.12)$$

□

Now let's consider gamma random variables. A finite sum of independent gamma random variables with the same scale parameter is a new gamma random variable with the same scale parameter, but the shape parameter is the sum of the addend shape parameters. In the next proof, we look at independent and indentially distributed gamma random variables.

Proposition 8.2. *If the u_i 's in the construction of our unistochastic matrix are independent $\Gamma(\alpha, \beta)$ random variables and $\frac{8}{\alpha} + 2 < n$, then*

$$E(\| M - I \|_F^2) < \frac{16n^5 \prod_{i=0}^7 (\alpha+i)}{\prod_{i=1}^8 [(n-1)\alpha-i]} \quad (8.13)$$

$$+ \frac{16n^3 \prod_{i=0}^3 (\alpha+i)}{\prod_{i=1}^4 [(n-1)\alpha-i]} \quad (8.14)$$

$$+ \frac{16n^5 (n-1) \prod_{i=0}^3 (\alpha+i)^2}{\prod_{i=1}^8 [(n-2)\alpha-i]}. \quad (8.15)$$

Proof. Previously we showed that almost surely

$$\| M - I \|_F^2 < \sum_{i=1}^n \frac{16}{(u_1^2 + u_2^2 + \dots + u_n^2 - u_i^2)^4} u_i^8 \quad (8.16)$$

$$+ \sum_{i=1}^n \frac{16}{(u_1^2 + u_2^2 + \dots + u_n^2 - u_i^2)^2} u_i^4 \quad (8.17)$$

$$+ \sum_{i \neq j} \frac{16}{(u_1^2 + u_2^2 + \dots + u_n^2 - u_i^2 - u_j^2)^4} u_i^4 u_j^4. \quad (8.18)$$

Using the Cauchy-Schwarz inequality and the fact that $0 < u_i$ almost surely for all i , we get

$$\| M - I \|_F^2 < \sum_{i=1}^n \frac{16n^4}{(u_1 + u_2 + \dots + u_n - u_i)^8} u_i^8 \quad (8.19)$$

$$+ \sum_{i=1}^n \frac{16n^2}{(u_1 + u_2 + \dots + u_n - u_i)^4} u_i^4 \quad (8.20)$$

$$+ \sum_{i \neq j} \frac{16n^4}{(u_1 + u_2 + \dots + u_n - u_i - u_j)^8} u_i^4 u_j^4. \quad (8.21)$$

If we take expected values, and take advantage of the independent and identically distributed u_i 's,

$$E(\| M - I \|_F^2) < E\left(\frac{16n^5}{(u_1+u_2+\dots+u_n-u_i)^8}\right)E(u_i^8) \quad (8.22)$$

$$+ E\left(\frac{16n^3}{(u_1+u_2+\dots+u_n-u_i)^4}\right)E(u_i^4) \quad (8.23)$$

$$+ E\left(\frac{16n^5(n-1)}{(u_1+u_2+\dots+u_n-u_i-u_j)^8}\right)E(u_i^4)E(u_j^4). \quad (8.24)$$

Since the u_i 's are gamma random variables and $\frac{8}{\alpha} + 2 < n$, we may replace the denominator sums with gamma random variables. Let γ_j denote a gamma random variable with parameters $j\alpha$, and β .

$$E(\| M - I \|_F^2) < \frac{16n^5 \prod_{i=0}^7 (\alpha+i)}{\prod_{i=1}^8 [(n-1)\alpha-i]} \quad (8.25)$$

$$+ \frac{16n^3 \prod_{i=0}^3 (\alpha+i)}{\prod_{i=1}^4 [(n-1)\alpha-i]} \quad (8.26)$$

$$+ \frac{16n^5(n-1) \prod_{i=0}^3 (\alpha+i)^2}{\prod_{i=1}^8 [(n-2)\alpha-i]}.$$

□

Next we will consider bistochastic matrices arising from random Householder matrices where

$$u_1 = u_2 = \dots = u_n, \text{ and } u_i \neq 0 \quad (8.28)$$

almost surely. It follows that almost surely

$$H = I - \frac{2}{n} \begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix}, M = \binom{n-4}{n} I + \begin{bmatrix} \frac{4}{n^2} & \dots & \frac{4}{n^2} \\ \vdots & \ddots & \vdots \\ \frac{4}{n^2} & \dots & \frac{4}{n^2} \end{bmatrix}. \quad (8.29)$$

For the previous matrix probability distributions described, we use random numbers to build matrices, then use the matrices to construct an empirical approximation to the probability distribution of the statistic under consideration. When $u_1 = u_2 = \dots = u_n$, $u_i \neq 0$ almost surely, the aspects of the matrix only depends on n , we know the precise distribution of all matrix statistics. If reasonable for the dynamical system under consideration, we could use M to establish critical values for hypothesis testing. So let's look at some aspects of such M 's, after a lemma.

Lemma 8.3. *If D_n and S_n are $n \times n$ symmetric matrices such that*

$$D_n = \begin{bmatrix} \alpha & \beta & \beta & \dots & \beta \\ \beta & \alpha & \beta & \dots & \beta \\ \beta & \beta & \alpha & \dots & \beta \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \beta & \beta & \beta & \dots & \alpha \end{bmatrix} \text{ and } S_n = \begin{bmatrix} \beta & \vec{\beta}^T \\ \vec{\beta} & D_{n-1} \end{bmatrix} \quad (8.30)$$

where $\vec{\beta}$ is the $n - 1$ vector with β for all entries, then

$$\det(D_n) = (\alpha - \beta)^{n-1}(\alpha + (n - 1)\beta) \text{ and} \quad (8.31)$$

$$\det(S_n) = (\alpha - \beta)^{n-1}\beta. \quad (8.32)$$

Proof by induction:. Assume true for n . Using the fact that interchanging any two rows or any two columns of a real matrix changes the sign of the determinant, we see that

$$\det(D_{n+1}) = \alpha \det(D_n) - \beta n \det(S_n) \text{ and} \quad (8.33)$$

$$\det(S_{n+1}) = \beta \det(D_n) - \beta n \det(S_n). \quad (8.34)$$

Using the induction hypothesis, these equations become

$$\det(D_{n+1}) = \alpha \left((\alpha - \beta)^{n-1}(\alpha + (n - 1)\beta) \right) - \beta n \left((\alpha - \beta)^{n-1}\beta \right) \text{ and} \quad (8.35)$$

$$\det(S_{n+1}) = \beta \left((\alpha - \beta)^{n-1}(\alpha + (n - 1)\beta) \right) - \beta n \left((\alpha - \beta)^{n-1}\beta \right). \quad (8.36)$$

Factoring out the $(\alpha - \beta)$ terms gives us the results. \square

Proposition 8.4. *If M is the $n \times n$ matrix matrix*

$$M = \left(\frac{n-4}{n} \right) I + \begin{bmatrix} \frac{4}{n^2} & \dots & \frac{4}{n^2} \\ \vdots & \ddots & \vdots \\ \frac{4}{n^2} & \dots & \frac{4}{n^2} \end{bmatrix}, \quad (8.37)$$

then $\det(M) = \left(\frac{n-4}{n} \right)^{n-1}$, $\text{trace}(M) = \frac{(n-2)^2}{n}$ and the Jordan canonical form of M is the diagonal matrix with entries $(1, \frac{n-4}{n}, \dots, \frac{n-4}{n})$.

Proof. The trace of M follows from the definition. The determinant follows from the previous lemma by setting $\alpha = \frac{(n-2)^2}{n^2}$, and $\beta = \frac{4}{n^2}$.

Now, the matrix is a symmetric real matrix; hence it is diagonalizable. The eigenvalues follow from the previous lemma by setting $\alpha = \frac{(n-2)^2}{n^2} - \lambda$, and $\beta = \frac{4}{n^2}$ to get the characteristic polynomial of M . \square

Since M is diagonalizable, if the Markov shift $((1/n, 1/n, \dots, 1/n), M)$ approximates a measure-preserving dynamical system that is mixing, the estimate of mixing rate is the rate at which

$$\left(\frac{n-4}{n}\right)^N \rightarrow 0 \text{ as } N \rightarrow \infty, \quad (8.38)$$

instead of the estimate given in [17, Chapter 4], $\binom{N}{n-1} \left(\frac{n-4}{n}\right)^{N-n+1} \rightarrow 0$ as $N \rightarrow \infty$.

9. Two Region Partitions

There are few instances of interest where one would use our method with two partition regions. We look at this special case as an example to help develop understanding. A potential application is equal ratio mixing of items with minimal consequences of poor mixing, such as combining blends of coffee. When combining two equal volumes of coffee, poor mixing would result in inconsistent taste. Only the most serious baristas would say that inconsistent cup-of-Joe flavor is worse than poorly mixing pharmaceuticals.

Say that our unit vector is

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}. \quad (9.1)$$

Since the vector has norm one, we may write its Householder matrix as

$$H = \begin{bmatrix} 1 - 2v_1^2 & \pm 2v_1\sqrt{1 - v_1^2} \\ \pm 2v_1\sqrt{1 - v_1^2} & 2v_1^2 - 1 \end{bmatrix}. \quad (9.2)$$

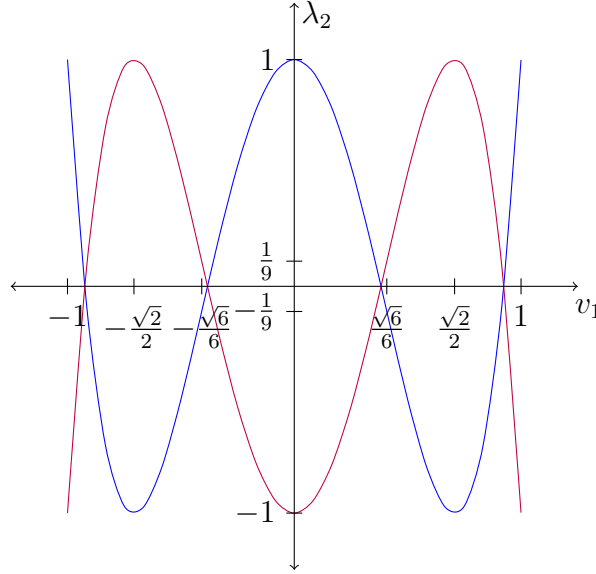
There are two possible doubly-stochastic matrices arising from our method,

$$\begin{bmatrix} (1 - 2v_1^2)^2 & 4v_1^2(1 - v_1^2) \\ 4v_1^2(1 - v_1^2) & (1 - 2v_1^2)^2 \end{bmatrix}, \text{ and } \begin{bmatrix} 4v_1^2(1 - v_1^2) & (1 - 2v_1^2)^2 \\ (1 - 2v_1^2)^2 & 4v_1^2(1 - v_1^2) \end{bmatrix}. \quad (9.3)$$

Whose characteristic polynomials are

$$(1 - \lambda)(8v_1^4 - 8v_1^2 + 1 - \lambda), \text{ and } (1 - \lambda)(-8v_1^4 + 8v_1^2 - 1 - \lambda). \quad (9.4)$$

The second largest eigenvalue depends on v_1 . Let's graph the relationship.



Graphs of the relationship between v_1 and λ_2 when $n = 2$

When $n = 2$, the relationship between v_1 and second largest eigenvalue defines a function from $[-1, 1]$ to $[-1, 1]$. In this situation, it is feasible to compute

$$P(|\hat{\lambda}_2 - 1| > k : \lambda_2 = 1) \text{ and } P(|\hat{\lambda}_2| > k : |\lambda_2| = 1). \quad (9.5)$$

If v_1 is a beta random variable with parameters α and β , then

$$E(\lambda_2) = \pm \frac{8\alpha(\alpha+1)(\alpha+2)(\alpha+3)}{(\alpha+\beta)(\alpha+\beta+1)(\alpha+\beta+2)(\alpha+\beta+3)} \quad (9.6)$$

$$\mp \frac{8\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} \pm 1, \quad (9.7)$$

where the sign of addends depends on the permutation matrix used.

References

- [1] G. BAND AND P. BOYLAND, *The Burau estimate for the entropy of a braid*, *Algebr. Geom. Topol.*, 7 (2007), pp. 1345–1378.

- [2] I. BENGTTSSON, Å. ERICSSON, M. KUŚ, W. TADEJ, AND K. ŻYCZKOWSKI, *Birkhoff's polytope and unistochastic matrices, $N = 3$ and $N = 4$* , Comm. Math. Phys., 259 (2005), pp. 307–324.
- [3] G. BIRKHOFF, *Three observations on linear algebra*, Univ. Nac. Tucumán. Revista A., 5 (1946), pp. 147–151.
- [4] P. BOYLAND AND J. HARRINGTON, *The entropy efficiency of point-push mapping classes on the punctured disk*, Arxiv preprint arXiv:1103.1829, (2011).
- [5] G. CASATI, G. COMPARIN, AND I. GUARNERI, *Decay of correlations in certain hyperbolic systems*, Phys. Rev. A, 26 (1982), pp. 717–719.
- [6] M. DELLNITZ, G. FROYLAND, AND S. SERTL, *On the isolated spectrum of the Perron-Frobenius operator*, Nonlinearity, 13 (2000), pp. 1171–1188.
- [7] J. DING AND A. ZHOU, *Finite approximations of Frobenius-Perron operators. A solution of Ulam's conjecture to multi-dimensional transformations*, Phys. D, 92 (1996), pp. 61–68.
- [8] G. FROYLAND, *Using Ulam's method to calculate entropy and other dynamical invariants*, Nonlinearity, 12 (1999), pp. 79–101.
- [9] —, *On Ulam approximation of the isolated spectrum and eigenfunctions of hyperbolic maps*, Discrete Contin. Dyn. Syst., 17 (2007), pp. 671–689 (electronic).
- [10] G. FROYLAND AND K. AIHARA, *Ulam formulae for random and forced systems*, Thinking, 1, p. 1.
- [11] G. H. GOLUB AND C. F. VAN LOAN, *Matrix computations*, Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, Baltimore, MD, third ed., 1996.
- [12] J. HARRINGTON, *Topological efficiency of stirring with obstacles*, PhD thesis, University of Florida, 2011.
- [13] F. Y. HUNT, *Unique ergodicity and the approximation of attractors and their invariant measures using Ulam's method*, Nonlinearity, 11 (1998), pp. 307–317.
- [14] T. Y. LI, *Finite approximation for the Frobenius-Perron operator. A solution to Ulam's conjecture*, J. Approximation Theory, 17 (1976), pp. 177–186.
- [15] C. MEYER, *Matrix analysis and applied linear algebra*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000. With 1 CD-ROM (Windows, Macintosh and UNIX) and a solutions manual (iv+171 pp.).

- [16] C. ROBERT AND G. CASELLA, *Introducing Monte Carlo Methods with R*, Use R!, Springer, 2009.
- [17] A. C. SMITH, *Using Ulam's method to test for mixing*, ProQuest LLC, Ann Arbor, MI, 2010. Thesis (Ph.D.)—University of Florida.
- [18] M. A. STREMLER, *Fluid mixing, chaotic advection, and microarray analysis*, in Analysis and control of mixing with an application to micro and macro flow processes, vol. 510 of CISM Courses and Lectures, SpringerWienNewYork, Vienna, 2009, pp. 323–337.
- [19] M. A. STREMLER, F. R. HASELTON, AND H. AREF, *Designing for chaos: applications of chaotic advection at the microscale*, Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 362 (2004), pp. 1019–1036.
- [20] J.-L. THIFFEAULT AND M. D. FINN, *Topology, braids and mixing in fluids*, Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 364 (2006), pp. 3251–3266.
- [21] S. M. ULAM, *Problems in modern mathematics*, Science Editions John Wiley & Sons, Inc., New York, 1964.